(54) Title: MULTI-CHANNEL SOUND REPRODUCTION SYSTEM FOR STEREOPHONIC SIGNALS

(57) Abstract: The invention concerns multi-channel reproduction of sound signals, more specifically derivation of a number of output sound signals from a pair of stereophonic signals such that each of these output signals can be reproduced via loudspeakers placed at the position of those phantom sources that would have been created by said stereophonic signals if they were provided to a pair of loudspeakers in a normal stereophonic set-up. The object of the invention is thus to replace said phantom sources by real physical sound sources, which would make the listening position in a room less critical as in a normal stereophonic set-up.

WO 01/62045 A1

MULTI-CHANNEL SOUND REPRODUCTION SYSTEM FOR STEREOPHONIC SIGNALS

5 TECHNICAL FIELD

The present invention relates generally to multi-channel sound reproduction via loudspeakers and more particularly to extraction of appropriate monophonic signal components from a normal stereophonic signal and providing each of these monophonic

10 signals to different loudspeakers in a multi-channel sound reproduction set-up.

DESCRIPTION OF PRIOR ART

15 A large number of multi-channel sound reproduction systems exist for use in large rooms such as cinemas or for use in smaller rooms such as normal domestic listening rooms. A specific class of such multi-channel systems utilises some kind of decoding means to translate the signals from two stereophonic sound tracks for instance on a motion picture film or on a gramophone record or compact disc for domestic use into a larger number of

20 signals, each of which is to be provided to separate loudspeakers placed at different positions in the listening room.

An example of a prior art system which is able to extract five loudspeaker signals for a left loudspeaker, a right loudspeaker and a centre loudspeaker placed midways between

25 the left and right loudspeaker in front of a listener and furthermore for a left and right loudspeaker placed behind the listener is disclosed in US-4024344 by Dolby et al. ("Centre Channel Derivation for Stereophonic Cinema Sound"). In this system, which is typically used in motion picture, theatre music and various moving sound effects are typically reproduced as stereophonic signals using the frontal left and right loudspeakers,

30 whereas speech is being reproduced as a monophonic sound signal from the frontal centre loudspeaker in order to obtain good speech intelligibility. The two-rear loudspeakers add reverberation to the reproduced sound field, giving the listener the impression of being surrounded by sound.

35 The extraction of the monophonic signal for the centre loudspeaker is in the above system based on determination of the correlation between the left and right stereophonic

signals. These signal components that are highly correlated with each other are extracted from the two channels, added and provided to the centre loudspeaker. There remains the "stereophonic part" of the signals which parts are reproduced via the front left and right loudspeakers as normal stereophonic signals.

A disadvantage with the above-mentioned system results from the fact, that in the centre channel signal the left and right channel signals are both represented with equal magnitude. Consequently, signal components in the left and right channel signals, differing only in that they have opposite phase but which are still highly correlated, disappear in the centre channel signal. The disappearance of such signal components from the centre channel leads to an unsatisfactory sound reproduction of that channel.

In order to avoid the above-mentioned disadvantage with the prior art system according to US-4024344, a system according to US-5426702 is suggested by Aarts. In this system a centre channel signal is derived from the left and right channel signals based on the determination of a direction vector which indicates the direction to the most powerful sound from origo in a coordinate system depicting the magnitude of the left signal along one axis and the magnitude of the right signal along the other axis. Based on this direction vector two weight factors are derived such that weighted right and left signals are added to form the centre channel signal. If the left and right signals are of opposite phase, the sign of the weight factors also differs, with the consequence that a centre channel signal is always obtained no matter whether the left and right channel signals are in phase or 180 degrees out of phase. This constant presence of a centre channel signal gives rise to a more satisfactory overall sound reproduction.

Apart from the fact that the prior art systems derive a purely monophonic signal to be provided to a centre loudspeaker, they still function to a large extent as a normal stereophonic loudspeaker system, i.e. the perceived sound images are the result of a perceptual combination in the brain of the listener of sound signal components originating from the left and right loudspeakers. If signal components from the left and right loudspeaker in such a system are either fully or at least partially correlated, these components will "melt together" in the brain of the listener into one spatially defined sound image, which will often be located somewhere on the line between the two loudspeakers. This perceived sound image is often termed a "phantom source", and it can be said that in stereophonic sound reproduction systems the formation of the overall perceived sound image basically relies on the formation of phantom sources. If either the

left or right channel signal is much stronger than the other, or there is a sufficient time delay between these signals, the phantom source will be located at one of the loudspeakers, i.e. either at the loudspeaker radiating the strongest signal or the loudspeaker leading in time relative to the other. Only in such cases there is a
5   coincidence between the phantom source and the actual physical sound source.

The fact that the prior art system at least to some extend relies on the formation of phantom sources in the creation of the perceived sound images gives rise to a number of problems. If a listener is placed directly in the symmetry plane between the left and right
10   loudspeakers and in a sufficient distance from the line between the loudspeakers, the listener as mentioned will perceive a sound image directly in the symmetry plane. If the listener moves for instance to the left relative to the symmetry plane, the magnitude of the signal received from the left loudspeaker will increase, and also the signal received from the left loudspeaker will arrive somewhat earlier than the signal from the right
15   loudspeaker. This will result in that the perceived sound image will move towards the left loudspeaker as the listener also moves in this direction. A sound element, which was intended to be located in the symmetry plane, will thus only be located in this plane, when the listener is also positioned herein. The optimal listening positions are thus confined to a narrow region around the symmetry plane. It would, however, be desirable
20   to extend the listening region to a large region of space, at least in front of the loudspeakers.

A localisation error not infrequently encountered in connection with the formation of phantom sources consists of a so-called elevation error, i.e. the phantom source, which
25   ideally should be perceived directly on the line between the left and right loudspeakers, and hence normally approximately at the level of the listener's ears, is actually being perceived above this level. Such elevation errors can be the result of the presence of small phase differences, which at a specific frequency correspond to similar minor time differences between the signals from the two loudspeakers at the position of the ears of
30   the listener. Such phase or time differences between two substantially equally powerful signals will produce a combfilter effect cancelling the sound signals at a discrete series of frequencies. Slight movements of the head of the listener will cause these cancellation frequencies to shift corresponding to the change in phase or arrival time of the left and right channel signals to the ears of the listener. The free field transfer functions of the
35   ears of human listeners however also exhibits such series of cancellation frequencies, where these cancellation frequencies depend on the elevation of a sound source relative

to the level of the listener's ears. If the cancellation frequencies produced by small phase or time differences between the left and right channel signals as received by the ears to some extent coincide with the cancellation frequencies of the free field transfer functions of the listener's ears, this coincidence can give rise to elevation errors.

Finally, it is a normal experience that phantom sources due for instance to small inevitable discrepancies between the amplitudes and phases between the left and right channel signals as received by the ears will be perceived as less spatially well-defined, i.e. more "diffuse", than the actual physical sounds sources, they are meant to represent.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a method and a device for spatial reproduction of a sound field based on stereophonic signals in a left and right channel, such as for instance stereophonic signals recorded on a normal compact disc, which reproduction does not suffer from the problems mentioned in the prior art section concerning a restricted optimal listening region around the symmetry plane between a left and right loudspeaker in a normal stereophonic reproduction system, so that listening can take place over most of the area of a normal listening room with equally satisfactory result.

It is a further object of the present invention to provide a method and a device of said kind, which does not suffer from unwanted elevation effects as described in the prior art section.

It is a further object of the present invention to provide a method and a device which reproduces the originally intended spatially well-defined sound sources as equally spatially well-defined sound images, no matter what position the listener is located in the listening room at least in front of the loudspeakers.

It is a further object of the present invention to provide a method and a device which can achieve these objects based on normal intensity stereophonic signals, and which method and device hence does not require specially encoded left and right channel signals to achieve said objects.

According to the invention these objects are achieved by replacing the phantom sources of a normal stereophonic reproduction system by a number of actual physical sound sources placed at the positions where said phantom sources would be located while listening to the normal stereophonic system from a ideal listening position substantially

5   located in the symmetry plane of the two stereophonic loudspeakers.

According to the invention a method is provided for converting two stereophonic (left and right channel) input signals L(t) and R(t) into N output signals according to the characterising clause of claim 1, where said method according to a preferred

10   embodiment of the invention comprises the following steps:

1. Based on the original left and right channel signals L(t) and R(t), intended for the left and right loudspeakers in a normal two-loudspeakers stereophonic reproduction system, and based on a comparison of each separate pair of left and right frequency

15   components (for instance provided by a fast Fourier transformation of said left and right signals) of these signals and on the application of a first specific set of requirements to the outcome of these comparisons extracting a first output signal as a linear combination of said left and right channel signals under the condition that the relationship between said left and right channel frequency components is such that

20   these would contribute to the formation of a first phantom source.

2. Providing a pair of first residual left and right channel signals, which pair does not contain those frequency components, which have been extracted in said step (1) and which would contribute to said first phantom source.

25

3. Based on the original left and right channel signals L(t) and R(t), and based in the same manner as above on a comparison of each separate frequency component of these signals and on the application of a second specific set of requirements to the outcome of these comparisons extracting a second output signal as a linear

30   combination of said first residual left and right channel signals under the condition that the relationship between said original left and right channel frequency components is such that these would contribute to the formation of a second phantom source located at a different position than said first phantom source.

4. Providing a pair of second residual left and right channel signals, which pair does not contain those frequency components, which have been extracted in said step (1) or (3) and which would contribute to said first and second phantom source.

5   5. Repeating the previous steps a sufficient number of times and each time with different sets of requirements to be able to extract N-2 output signals corresponding to N-2 phantom sources, which could be formed by the original left and right channel signals L(t) and R(t).

10 6. Providing a pair of final residual left and right channel signals, which pair does not contain those frequency components, which would contribute to any of said first, second, etc. phantom sources.

7. Providing said first, second, etc. output signals to electroacoustic transducers, e.g.
15 loudspeakers, the position of each of these loudspeakers corresponding to the particular set of requirements utilised at the extraction of the output signal for that particular loudspeaker.

8. Providing said final residual left channel signal to an electroacoustic transducer, e.g.
20 a loudspeaker, placed to the left of all other N-1 loudspeakers and providing said final right channel signal to an electroacoustic transducer, e.g. a loudspeaker, placed to the right of all the other N-1 loudspeakers.

As an alternative to the above step (3) of the method according to the invention said
25 comparison and application of the specific set of requirements could be carried out on the pair of first residual left and right channel signals provided in step (2) above instead of on the original left and right channel signals. It is advantageous that the procedure described in step (3) is applied, but in a practical implementation it may be necessary or desirable to apply said alternative.
30

According to the invention there is furthermore provided a device according to the characterising clause of claim 16 for carrying out said method, where said device comprises N-2 means for extracting said output signals corresponding to said phantom sources, where each of said N-2 means furthermore provides said pairs of residual left
35 and right channel signals which does not contain any of - or according to a second embodiment only a fraction of - those signal components, that would have contributed to

said phantom sources, which pair of residual signals are provided to succeeding means for extraction of the remaining output signals.

The extractions of said output signals from the left and right input signals – or from the
5   corresponding residual signals – is according to the invention based on a running comparison, i.e. a comparison as a function of time, of the degree of linear dependency between each of said pairs of separate frequency components of the two input signals. A measure of the degree of linear dependency between left and right signals is thus according to the invention based on a running cross correlation analysis of left and right
10   signal pairs and a succeeding determination of the coherence function, which is a number between 0 and 1, where the value 1 is obtained when the left and right signals are fully correlated and the value 0 is obtained when the left and right signals are fully uncorrelated.

15   According to the invention the criterion for extraction of a output signal to be provided to one of said N-2 loudspeakers positioned between the left and right loudspeakers is that the coherence function should have a value close to 1, preferably between 0.8 and 1, although other intervals may also be chosen. If it is found that certain left and right signal elements fulfil said coherence criterion, those elements could have contributed to the
20   formation of a phantom source in a normal left and right channel stereophonic system, and will thus according to the invention be represented by an actual physical sound source, i.e. one of the N-2 loudspeakers placed between the outermost left and right loudspeakers. The signal to be provided to this loudspeaker is according to the invention being obtained by a linear combination of the corresponding left and right input signals to
25   that particular processing block, in which the extraction of that particular output signal takes place.

Which one of these N-2 loudspeakers actually should be provided with the extracted signal could on principle be determined based on either a comparison – for each pair of
30   frequency components - of the magnitudes of the left and right signals or on a comparison of the relative phase (or time delay) between these frequency components. It is also possible to use combinations of magnitude and phase (or time) differences for extracting a measure for the lateralisation of the phantom source and hence for the appropriate location of a corresponding sound source.
35

More details about the determination of the coherence function and the magnitudes and relative phase of the left and right signals will be given below, together with a general description of the manner in which the output signals are being extracted and the residual left and right signals obtained.

5

As is apparent from the foregoing description the system according to the invention can be said to replace the phantom sources obtained in a normal stereophonic system with a corresponding number of real physical sound sources. In normal stereophonic systems phantom sources will only be perceived, if correlated signal components are found in the

10   left and right channels (see for instance: Jens Blauert, "Spatial Hearing", Section 3.1.). The perceived position of the phantom source will depend on both the amplitude difference and the phase difference (or time difference) between the correlated signal components, this dependency being generally a function of frequency, in the left channel relative to the right channel. If for instance the signal in the right channel is louder

15   compared to the signal in the left channel, a phantom source will be perceived at a position to the right of the symmetry plane between the two loudspeakers. Analogous to this situation if the right channel is delayed compared to the left channel, the phantom source will be perceived to the left of said symmetry plane. A time delay corresponds to a linear phase difference, i.e. a phase difference, which is proportional to the frequency.

20

A majority of normal stereophonic recordings of for instance music is based on the technique called "intensity stereophony", i.e. amplitude differences between the two channels are being used to create phantom sources. As a measure of the degree of correlation between the left and right signal, the coherence function $\gamma(f)$ can be used. The

25   coherence function is a real number between 0 and 1 indicating the fraction of power in the correlated part of the signals compared to the total signal power, when considering two signals, for instance the left signal $L(t)$ and the right signal $R(t)$ in a normal stereo system. The coherence is 1 when the two signals are fully correlated at that frequency, i.e. when the L and R signals are linear functions of each other, such as identical signals

30   or one signal and a delayed and/or scaled version of this signal. A value of the coherence function of 0 indicates totally uncorrelated signals. Equation (1) gives the coherence function $\gamma(f)$ at the frequency f, obtained using calculated values of the cross spectrum $G_{LR}(f)$ and the two auto spectra G_{LL} and G_{RR} based on the spectra $L(f)$ and $R(f)$ obtained by FFT analysis of the original pair of signals $L(t)$ and $R(t)$. For more information

35   about the coherence function see for instance Julius S. Bendat and Allan G. Piersol:

"Engineering Applications of Correlation and Spectral Analysis", published by "Wiley-Interscience", ISBN 0-471-57055-9.

$$\gamma(f) = \frac{|G_{LR}(f)|}{\sqrt{G_{LL}(f) \cdot G_{RR}(f)}}$$

(1)

Both amplitude difference and time delay between the left and right signals are crucial when predicting the position of a phantom source, see above-mentioned reference to J. Blauert. Using equation (2), the amplitude difference amp(f) can be calculated based on the two auto spectra G_{LL} and G_{RR}:

$$amp(f) = \sqrt{\frac{G_{LL}(f)}{G_{RR}(f)}}$$

(2)

A pure time delay corresponds to a linear phase, i.e. a linear dependency between phase shift and frequency. Equation (3) gives the phase shift phase(f) calculated as the angle of the complex valued cross spectrum of left and right signals G_{LR}:

$$phase(f) = angle(G_{LR}(f))$$

(3)

The group delay grd(f) is a measure of the delay of a narrowband signal centred on a frequency f. Equation (4) gives the group delay grd(f) calculated from the unwrapped/continuos phase, continuos-phase(f). The continuous phase can be found from equation (3) by adding or subtracting an appropriate number of 2π radians at different frequencies, so that the phase becomes a continuous function of frequency:

$$grd(f) = -\frac{d(continuous\_phase(f))}{d\omega}$$

(4)

The group delay according to equation (4) is obtained in a number of samples, and a division by the sampling frequency f_s hence gives the group delay $\tau(f)$ expressed in seconds, see equation (5):

5

$$\tau(f) = \frac{-d(continous\_phase(f))}{2\pi \cdot df}$$

(5)

10 Certain requirements must be fulfilled before a part of the left and right signals are extracted and provided to a specific loudspeaker. These requirements comprise upper and lower limits on the amplitude difference between left and right signals, limits on group delay between these signals and as mentioned previously a minimum value of the coherence function. These three requirements together ensure that a phantom source

15 was intended to be formed in the vicinity of a given one of the loudspeakers.

Enforcement of the limits can be carried out very sharply as in said first embodiment of the invention or smoothly as in said second embodiment of the invention. A sharp enforcement is obtained by requiring that the value of the coherence function should be

20 at least 0.8 for a signal to be extracted for a specific loudspeaker. A smooth enforcement would be obtained by providing a highly attenuated signal to the particular loudspeaker at a coherence value of for instance 0.7 and letting the signal level increase gradually up to coherence values above 0.9.

25 According to the first embodiment of the invention the sharp limits are used, i.e. the total left and right signals components at a given frequency are extracted after suitable combination hereof and provided as an output signal to the particular loudspeaker.

Different sets of requirements are to be met in extracting the signals to be provided to the

30 different N-2 loudspeakers positioned between the left and right loudspeaker. A requirement R comprises three parameters: the minimum value of the coherence function, the range of the amplitude difference (dB) between left and right signal and the range of the group delay (ms) or phase difference (degrees) between left and right signals.

35

For a specific embodiment of the system according to the invention to be described in detail in the detailed description of the invention N equals 5, and thus three loudspeakers – centre-left, centre and centre-right – are placed substantially equidistantly between the left and right loudspeakers. For this particular embodiment the different sets of requirements could for instance be the following, although other requirements and/or specific values would also be conceivable:

Centre channel: coherence > 0.8, amplitude difference <= +/- 2dB, group delay difference <= +/- 2ms (or phase <= +/- 20 degrees)

Centre-left channel: coherence > 0.8, amplitude difference +2dB to +6dB, group delay difference <= +/- 2ms (or phase <= +/- 20 degrees)

Centre-right channel: coherence > 0.8, amplitude difference -2dB to -6dB, group delay difference <= +/- 2ms. (or phase <= +/- 20 degrees)

In the above set of requirements only amplitude differences are used to decide between the different loudspeakers. It is as mentioned previously also possible to base the choice between the loudspeakers on group delay differences (or phase differences which are related to group delay differences at a specific frequency) or on combinations of amplitude- and group delay(phase) differences. It should be emphasised that the invention is not limited to the utilisation of amplitude differences for the choice between the different loudspeakers, although a choice based on amplitude differences may be advantageous, because the normal way of producing stereophonic signals (so-called intensity stereophony), i.e. left/right channel signals, to be recorded for instance on normal compact discs is to control the lateralisation of the created phantom sources by manipulating the relative amplitudes (levels) of different output sound recordings in an electronic mixing console. Creation of phantom sources by manipulating relative group delays of output signals is normally not used.

According to the invention a fourth requirement is set up in order to handle the special case of left and right signals being in anti-phase, i.e. 180 degrees out of phase. If the left and right channel signals are 180 degree out of phase the corresponding group delay is still 0 ms. Consequently, two otherwise identical signals in the left and right channels but 180 degrees out of phase will fulfil the above three requirements for a signal to be extracted and provided to the centre channel. As mentioned the extracted output signal is formed as a linear combination of left and right channel signals. According to a

preferred embodiment of the invention this linear combination consists of the sum of the left and right channel signals and in the case of 180 degrees phase shifted left and right signals the extracted output signal will thus be equal to zero. As those signal components which are being extracted from the left and right channel signals equal the

5 total left and right signals respectively in this thus both the extracted monophonic signal and the residual left and right signals will be equal to zero and consequently no sound will be radiated from any of the N loudspeakers. This is a clearly unwanted situation, and a fourth requirement to avoid this situation from occurring could be that the phase difference should be kept within of instance +/- 170 degrees at any times in order to

10 allow an extraction of output signals.

As an alternative to this fourth requirement the previously mentioned requirement on the group delay difference <= +/- 2ms could be replaced by a limitation of the allowable phase difference between the left and right signals, so that for instance only signal

15 elements with relative phase differences between left and right signal <= +/- 20 degrees are used for the extraction of output signals. The above-mentioned fourth requirement would in this case be unnecessary.

According to the second embodiment of the invention the extraction is still based on

20 specific sets of requirements for the coherence function, the amplitude difference and the phase difference corresponding to each of the phantom sources, which in this case generally are only partly replaced by physical sound sources. In the second embodiment of the invention, however, the fraction of each frequency component to be extracted from the specific input signals is obtained by multiplying these frequency components with a

25 filter function H(z) which is a product of continuos functions the parameters of which are chosen according to the specific sets of requirements, e.g. Gaussian functions (normal distribution density function) of the values of the square of the coherence function, the amplitude difference and the phase difference, where the parameters of these three Gaussian functions (normal distribution density function) (means and variances)

30 correspond to sets of requirements as for instance those used in the first embodiment. Thus, for instance if a signal to the center loudspeaker is to be extracted, the mean value of the three Gausian functions will be 1 (coherence), 0 (amplitude difference) and 0 (phase difference), and the variances will be suitably chosen, so that the product of these three Gaussian functions (normal distribution density function) will only be substantially

35 equal to unity for those signal components that correspond to the particular phantom source. which is to be replaced entirely by a physical sound source. The value of the

filter function H(z) can thus be anywhere between 0 and 1, yielding a more smooth enforcement of the requirements for extraction of monophonic output signals than obtained according to the first embodiment.

5   According to a third embodiment of the invention it is possible to combine said sharp enforcement of the requirements for extraction of monophonic output signals according to the first embodiment and said smooth enforcement according to the second embodiment described above. This can for instance be done by replacement of said filter function H(z) according to the first or second embodiment with a new filter function H(z)

10   formed as a product of a logical function H1(z;p) with output values of 1 or substantially 0 according to whether the parameters p, which may be the coherence function, the amplitude difference, the phase and/or group delay difference, belongs to the corresponding target intervals according to the first embodiment, and a function H2(z;q) which according to the second embodiment is a product of continuous functions, where q

15   denotes the remaining parameters not contained in said function H1.

BRIEF DESCRIPTION OF THE DRAWINGS

20   The invention will now be described in more detail with reference to the accompanying drawings, in which
     Figure 1 is a view of a normal stereophonic loudspeaker set-up also depicting the formation of a phantom source;
     Figure 2 is a view of a five loudspeaker set-up in a system according to the present
25   invention;
     Figure 3 is an embodiment of the system according to the present invention utilising three of the processing blocks shown in fig. 3 and N=5, i.e. a total of five loudspeakers;
     Figure 4 is a block diagram of a single processing block in the system according the invention;
30   Figure 5 is a detailed block diagram of one of the processing blocks shown in fig. 3 and 4 according to the first embodiment of the invention;
     Figure 6(a) is a detailed block diagram of the final stages of the analysis part of the system according to the first embodiment of the invention;
     Figure 6(b) is a detailed block diagram of the final stages of the analysis part of the
35   system according to the second embodiment of the invention; and

Figure 7 is a block diagram of the preferred arrangement of the whole system according to the invention.

5   DETAILED DESCRIPTION OF THE INVENTION

In the following a detailed description of one specific embodiment of the invention is given. In this embodiment N = 5, i.e. a total of five loudspeakers are used and these loudspeakers are placed in a line in front of a listening area, although the loudspeakers

10   could also have been placed along for instance an arc in front of the listening area.

With reference to fig. 1, a normal stereophonic loudspeaker set-up is shown. An actual physical sound source located midways between the two loudspeakers is in this set-up being simulated with the aid of two highly correlated electrical signals L(t) and R(t) fed to

15   the loudspeakers. These signals give rise for a listener located substantially midways between the loudspeakers – approximately at position A on the figure – to the creation of a phantom source 16, i.e. a sound source is perceived midways between the loudspeakers as indicated by the area PS in the figure. If, however, the listener is located off the symmetry plane between the two loudspeakers 11, 12 and specifically to the left

20   of the symmetry plane (for instance in seat 14), the perceived sound image is no longer located at PS as intended but is shifted more or less to the left as indicated at 17 by the area B in the figure. The overall perceived sound image thus depends on the position of the listener, and the "correct" perception of a sound source at PS is thus only obtained in a narrow region around A in the figure.

25

Figure 2 shows one embodiment of a system according to the present invention utilising five loudspeakers 21, 22, 23, 24, 25 placed in front of a row of seats 26, 27, 28 in a listening room. Thus N = 5 in this embodiment. In the system according to the invention a physical sound source midways between the extreme left and right loudspeakers 21 and

30   25 is not simulated by a phantom source midways between these loudspeakers but by a physical sound signal radiated by the centre loudspeaker 23. This means, that a listener will perceive the sound as originating from the centre loudspeaker 23 no matter where he is located, at least in the whole listening area in front of the loudspeakers. Hence correct spatial reproduction of a given original sound source is being preserved by the system

35   according to the invention, no matter what listening position the listener actually chooses.

The correct spatial characteristics of the perceived sound image are also preserved, if the listener moves around ir front of the loudspeakers.

Figure 3 shows an embodiment of the system according to the present invention utilising

5   three processing blocks 32, 33, 34 and five loudspeakers 35, 36, 37, 38, 39. A normal intensity stereophonic signal L, R is provided from a stereophonic source 31, exemplified by a CP-player, to the first processing block 32. This processing block 32 extracts in a manner to be described in detail in the following an output centre channel signal c_1, which is being provided to the centre loudspeaker 37. The output signal c_1 is in

10   processing block 32 being removed from the left and right signals L and R in a manner to be described in detail in connection with the description of fig. 4, 5 and 6(a) and 6(b), and two residual left and right signals L' and R' are being forwarded as input signals to the next processing block 33. This processing block 33 extracts in an analogous manner as block 32 a second output signal c_2, which is being provided to a loudspeaker placed

15   midways between the left loudspeaker 35 and the overall centre loudspeaker 37. The second output signal c_2 is removed from the signals L' and R' in a manner analogous to the procedure in the preceding block 32, and two new output signals L" and R" are being obtained and forwarded as new input signals to the succeeding processing block 34. An analogous process is again carried out in this block, extracting a third output signal c_3,

20   which is provided to a loudspeaker 38 placed midways between the right loudspeaker 39 and the overall centre loudspeaker 37. Finally, two output signals L'" and R'" are "left over", which are being provided to the left loudspeaker 35 and the right loudspeaker 39, and which signals constitute an intensity stereophonic signal pair, from which pair the signals to the three intermediate loudspeakers 36, 37, 38 have been removed.

25

The basic structure of these processing blocks 32, 33, 34 is shown by the block diagram in fig. 4. This figure shows how a fraction of the separate frequency components of the left and right channel 41, 42, i.e. that fraction of frequency components which fulfils the specific set of requirements for the particular output signal 411, which is to be extracted

30   by that particular processing block 40, can be separated and provided as an output signal c1, c2, c3 ... to one new channel, e.g. the centre channel. The filter H(z) 43, which according to this embodiment of the invention in principle can only take on the two values 1 or 0 at any given frequency, is used to filter both left and right channel signals 41, 42, and thereby to isolate those parts of the left and right channel signals, which fulfils the

35   particular requirements for that output signal, which is to be provided to that channel and removed from the left and right input signals in order to produce the residual left and right

channel signals L' and R' respectively. It should be noted that the filter 43 used for the left channel is similar to the filter used for the right channel.

The rationale behind this choice of filter is that if the requirements mentioned above are
5   fulfilled, the original stereophonic signal might have been produced by panning, i.e. splitting an output signal up into two parts, which are provided to the left and right channels separately. In intensity stereophony panning consists of splitting an output signal up into two signals with an appropriate amplitude (intensity) difference between the two signals and adjusting this amplitude difference, so that it corresponds to the
10   desired lateral position of the finally created phantom source. Hence separating these two parts of the left and right channel signals by applying the same filter 43 possibly weighted by an appropriate gain factor to the frequency components of the two signals allows the reconstruction of the original single signal by adding these two parts.

15   The frequency components of the output signals of the filters H(z) 43 are added in an addition means 45 to produce an output signal 48 and a gain 49 and post delay 410 is applied to this signal to obtain the desired output signal 411. The gain 49 can be used to adjust the output level of the signal radiated from the particular loudspeaker, to which the signal 411 is being provided, in order for instance to preserve total radiated power. The
20   post delay 410 will be explained in the following.

As mentioned, the parts of the left and right channel signals L and R which are extracted and provided as an output signal to the particular channel should be removed from the left and right channels, leaving the residual left and right signals L' and R' respectively.
25   This is done by subtracting in subtraction means 44 the output signals from H(z) 43 from delayed versions, delayed in two delay means 48, of the left and right channel signals. This delay is introduced to compensate for the delay of H(z) 43, which should ideally be a linear phase filter, i.e. exhibit a frequency independent delay.

30   If only three loudspeakers are used, i.e. N=3, these three loudspeakers are connected to the outputs 46 (left loudspeaker), 47 (right loudspeaker) and 411 (centre loudspeaker) and the post delay 410 is in this case set to 0. If as in the present embodiment five loudspeakers are used, i.e. N=5, the post delays 410 in each of the three processing blocks shown in figure 3 are adjusted to compensate for the processing delays in the
35   various blocks 32, 33, 34, so that the signals from all five loudspeakers 35, 36, 37, 38 and 39 are radiated synchronously.

The strategy of the total system as shown in figure 3 is that each processing block 32, 33, 34 takes those parts of the left and right channel signals, which fulfil the requirements set up for each loudspeaker, and then passes the remaining parts (the residual left and

5   right signals) on the next processing block in the chain. The residual left and right signals, which remain after the processing in the last of the preceding blocks 34 have been carried out, are then provided to the left 35 and right 39 loudspeakers respectively. This ensures that if no parts of the left and right channel signals fulfils the requirements set up for any of the intermediate loudspeakers 36, 37, 38, then the signals are

10   reproduced only by the outermost left and right loudspeakers 35, 39 as ordinary stereophonic reproduction.

H(z) 43 is calculated independently at different frequencies or in a number of different frequency bands. One way of calculating H(z) consists of evaluating the logical

15   expression of a form given by equation (6):

$$H(z)=(\gamma(z)>0.8)AND(-2dB<amp(z)<2dB)AND(-20^{\circ}<phase(z)<20^{\circ})$$

(6)

20   This equation returns the value 1 if all of the requirements are met at a given frequency z(rad/sample) and otherwise it returns the value 0. Consequently, the gain of H(z) at any frequency is either 1 or 0. This can lead to numerical problems when implementing H(z), and the value 0 can therefore be substituted by a finite attenuation, e.g. 0.001, which makes the numerical problems less pronounced. Other problems may arise when a

25   signal parameter (coherence, amplitude difference, group delay/phase) corresponds to a limit of one of the requirements, e.g. amp(z) = -2dB. In this situation the slightest change of parameter value can make H(z) shift from 0 to 1 or from 1 to 0. To avoid this problem a hysteresis can be implemented by changing the limits once a requirement has been met, e.g. to (-2.5 dB < amp(z) < 2.5 dB). In this case the value of amp(z) needs to

30   change more than 0.5 dB before it can make H(z) shift back to 0.

Since each gain of H(z) at different frequencies are calculated independently, very sharp transitions can be found in H(z) when viewing across frequencies. A smoothing 519 might then be applied to the target of H(z) before implementing H(z), e.g. a Gaussian

35   function (normal distribution density function) with frequency dependent width (e.g. 1/3 octave).

Figure 5 contains a detailed block diagram of the processing block shown in figure 4. The upper part of figure 5 (reference numerals 51 to 521) and figure 6(a) shows the determination of the function H(z) based on left and right input signals 51, 52 according

5   to the first embodiment of the invention, and the lower part of figure 5 (reference numerals 522 to 534) corresponds to figure 4 except for the fact that in figure 5 fast convolution (see Oppenheim and Schafer: "Descrete-time-signal-processing", Prentice Hall, 1989, ISBN 0-13-216771-9) is employed to perform convolution by H(z). In fast convolution the time domain signals l(n) and r(n) are fast Fourier transformed by means

10   524, a multiplication with H(z) is carried out by multiplication means 526 and an inverse fast Fourier transform 527 is carried out on the output signals from the two multiplication means 526.

The determination of H(z) carried out in the upper part of figure 5 and in figure 6(a) are

15   based on block operations, e.g. 512 samples at a time. These samples are isolated using time windows 53. After a transformation to the frequency domain has been performed by FFT means 54, three quantities are calculated by means 55, 56 and 57: the instantaneous autospectra G_{ll} and G_{rr} are calculated in 55 and 56 respectively and the instantaneous crossspectrum G_{lr} is calculated in 57. These instantaneous spectra are

20   then turned into a real estimate of these spectra by the application of low pass filtration in each of the filters 58 respectively, one frequency at a time. This is done in this embodiment of the invention using first order IIR filters for each frequency. The foregoing equations (1), (2) and (3) are then used to calculate the desired parameters, i.e. the phase difference is calculated in 511, the coherence function is calculated in 512 and the

25   amplitude difference is calculated in 513. After these calculations have been carried out, the resulting parameter values are compared with the set of requirements corresponding to the particular output signal, which it is desired to derive, and this is done by comparing in blocks 514, 515 and 516 the output signals from the means 512, 511 and 513 with the specific parameter target ranges corresponding to the particular output signal ($c1$, $c2$, $c3$

30   ...), which is to be extracted as exemplified by the parameter intervals for the phase, the coherence and the amplitude difference shown in fig. 6(a), so that according to the result of these comparisons three logical values 1 or 0 are obtained and provided as input signals to the logical AND block 517, which will provide an output value of 1, if all three requirements are fulfilled and 0, if one or more of these requirements is/are not fulfilled.

35   After the logical values of H(z) are obtained, a finite attenuation is as mentioned previously substituted for H(z) = 0 (block 518). After each block of 512 samples have

been processed, a new filter H(z) is being determined. When a new filter H(z) suddenly changes, this change can result in a "click" in the output signal. In order to avoid this problem a slew rate limiter 519 at each frequency is inserted after block 518. This means that the gain at any frequency is not allowed to change more than for instance +/- 0.08

5   dB/block.

The lower part of figure 5 (522 through 534) is the processing part of the system according to the invention while the upper part of figure 5 (51 through 521) is the analysis part of the system. If the complete system contains more than 1 processing block

10   (corresponding to N = 3, i.e. three loudspeakers) such as the system according to the present embodiment of the invention as shown in figure 3 (corresponding to N = 5, i.e. 5 loudspeakers and three processing blocks), two possible configurations of the series of processing blocks would be possible. According to the first of these configurations the two input terminals 51, 52 of the analysis part of the block are connected to the

15   corresponding two input terminals 522 and 523 respectively of the processing part of the block. This means that the input signals to the analysis parts of the first block would be the original left and right channel signals L and R, that the input signals to the analysis part of the next block would be the residual left and right channel signals L' and R' and so on. If, during the production of the original stereo signals, an output signal is being

20   rapidly panned between the left and right channel for instance simulating a rapid shift of the position of a sound source between for instance the centre loudspeaker 37 and the loudspeaker 38 to the right of this, there will initially correctly be extracted an output signal for the centre loudspeaker 37and finally also correctly an output signal for the loudspeaker 38 to the right of the centre loudspeaker. Due to the inevitable processing

25   delay in the analysis parts of the three blocks a certain time interval will elapse between the extraction of the first of said output signals and the second one of these, and in the meantime no output signals will be extracted by any of the blocks of the system, and the original left and right channel signals will move all through the cascaded processing parts of the system so that the final residual left and right channel signals L''' and R''' will be

30   equal to L and R respectively and hence the original stereophonic signals will in this intermediate time interval be – erroneously - played back by the left and right loudspeakers. Thus, a shift of the perceived sound image from a position at the centre loudspeaker 37 directly to the position of the loudspeaker 38 to the right thereof will not be obtained but rather a transition from a spatially well-defined sound image at the centre

35   loudspeaker 37 followed by a "broadening" or "smearing out" of the perceived sound

image and finally followed by the formation of a spatially well-defined sound image at loudspeaker 38.

In order to avoid these unwanted perceptual phenomena the input terminals 714, 715;
716, 717; 718, 719 to all of the analysis parts of the blocks should be connected in parallel and connected to the original left and right channel signals L and R. These considerations lead to the following preferred arrangement of the present invention shown in figure 7. In this figure the input terminals 714, 715; 716, 717; 718, 719 to the three analysis parts 73, 74, 75 are all connected to the original left and right channel signals, whereas the three processing blocks 76, 77, 78 extracting output signals for the centre loudspeaker 711, the loudspeaker 710 to the left of the centre loudspeaker 711 and the loudspeaker 712 to the right of the centre loudspeaker 711 are coupled in series as already shown in figure 3.

According to the second embodiment of the invention as shown in figure 6(b) the phase difference, phase, is at 61 provided to a means 64 for calculation of the exponent of the corresponding Gaussian function (normal distribution density function), which Gaussian function (normal distribution density function) in the case shown in figure 6(b) corresponds to the extraction of signal components corresponding to a phantom source placed directly midways between the outermost left and right loudspeakers, and hence the mean of this Gausian function (normal distribution density function) is 0. Similarly, the squared coherence function is at 62 provided to a means 65 for calculation of the exponent of the second one of said three Gaussian functions (normal distribution density function) and the amplitude difference is at 63 provided to means 66 for calculating the exponent of the third one of said Gaussian functions (normal distribution density function). The three Gaussian functions (normal distribution density function) are hereafter calculated in three identical means 67, the output of each of these being provided to a multiplication means 68, which via a succeeding slew rate limiter 69 and smoothing 610 provides the final filter function H(z), 611, the value of which will be equal to unity for those frequency components which correspond exactly to a phantom source midways between said outermost left and right loudspeakers, and less than unity for frequency components corresponding to a phantom source created somewhat either to the left or to the right of the center loudspeaker or for frequency components, which do not correspond to any phantom source, because the corresponding coherence function differs significantly from unity.

The signal provided by the slew rate limiter 69 and succeeding smoothing 610 is hereafter used as a weigthing function, and provided to the multiplication means 526 shown in figure 5.

5   Although various embodiments of the present invention have been shown and described in the preceding parts of the detailed description, it is understood that a person skilled in the art may conceive other embodiments of the invention without departing from the scope of the invention as defined by the following claims.

10

CLAIMS:

1. Method for converting two input signals L(t) and R(t) constituting the signals in the left and right channel of a stereophonic signal into N output signals constituting N output channels, where N > 2, characterised in that said output signals (c1, c2, c3 ...) are being extracted from said two input signals based on the instantaneous degree of linear dependency between said two signals and utilising sets of requirements concerning characteristic differences between said two input signals, said requirements being specific for each of said output signals (c1, c2, c3...).

2. Method according to claim 1, characterised in that said method comprises the following steps:

(A) Based on the original left and right channel signals L(t) and R(t), intended for the left and right transducers in a normal two-loudspeakers stereophonic reproduction system, and based on a comparison of each separate pair of left and right frequency components (for instance provided by a fast Fourier transformation of said left and right signals) of these signals and on the application of a first specific set of requirements to the outcome of these comparisons extracting a first output signal (c1) as a linear combination of said left and right channel signals under the condition that the relationship between said left and right channel signal components is such that these would contribute to the formation of a first phantom source.

(B) Providing a pair of first residual left and right channel signals (L', R'), which pair does not contain those scaled versions of frequency components. which has been extracted in the preceding step (A).

(C) Based on the original left and right channel signals L(t) and R(t), and based in the same manner as above on a comparison of each separate frequency component of these signals and on the application of a second specific set of requirements to the outcome of these comparisons extracting a second output signal (c2) as a linear combination of said residual left and right channel signals under the condition that the relationship between said original left and right channel signal components is such that these would contribute to the formation of a second phantom source located at a different position than said first phantom source.

(D) Providing a pair of second residual left and right channel signals (L", R"), which pair does not contain those scaled versions of frequency components. which have been extracted in the preceding steps (A and C).

(E) Repeating the previous steps a sufficient number of times and each time with different sets of requirements to be able to extract a maximum of N-2 output signals (c3, c4 ...) corresponding to N-2 phantom sources, which could be formed by the original left and right channel signals L(t) and R(t).

5   (F) Providing a pair of final residual left and right channel signals (L''', R'''), which pair does not contain those scaled versions of frequency components, which have been extracted in any of the preceding steps.

(G) Providing said first, second, etc. output signal (c1, c2, c3 ...) to electroacoustic transducers (35, 36, 37, 38, 39), the position of each of these transducers corresponding

10   to the particular set of requirements utilised at the extraction of the output signal (c1, c2, c3 ...) for that particular transducer.

(H) Providing said final residual left channel signal (L''') to an electroacoustic transducer placed to the left of all other N-1 transducers and providing said final residual right channel signal (R''') to an electroacoustic transducer placed to the right of all the

15   other N-1 transducers.

3. Method according to claim 2, <u>characterised in</u> that said comparison of the original left channel signal L(t) and the original right channel signal R(t) comprises the determination at each frequency component of the coherence function (γ) of said original signals L(t)

20   and R(t), the amplitude difference (amp) between said original signals L(t) and R(t) and the phase (or group delay) difference (phase or τ) between said original signals L(t) and R(t).

4. Method according to claim 3, <u>characterised in</u> that said coherence function (γ), said

25   amplitude difference (amp) and said phase (or group delay) difference are functions of frequency and are calculated on the basis of the crossspectrum $G_{LR}(f)$ and the two autospectra $G_{LL}(f)$ and $G_{RR}(f)$ according to the following equations:

30

$$\gamma(f) = \frac{|G_{LR}(f)|}{\sqrt{G_{LL}(f) \cdot G_{RR}(f)}} \qquad (1)$$

35

$$amp(f) = \sqrt{\frac{\overline{G_{LL}(f)}}{\overline{G_{RR}(f)}}} \qquad\qquad (2)$$

5

$$phase(f) = angle(G_{LR}(f)) \qquad\qquad (3)$$

10

$$\tau(f) = \frac{-d(continous\_phase(f))}{2\pi \cdot df} \qquad\qquad (5)$$

5. Method according to any of the preceding claims, <u>characterised in</u> that said sets of
15   requirements each comprise a target interval of said coherence function (γ), a target
interval of said amplitude difference (amp) and a target interval of said phase- or group
delay difference (phase, τ), which target intervals may be functions of the frequency.

6. Method according to any of the preceding claims, <u>characterised in</u> that said extraction
20   of output signals (c1, c2, c3...) is based on a comparison at each frequency component
of said coherence function (γ), said amplitude difference (amp) and said phase or group
delay difference (phase, τ) with the respective one of said target intervals, such that a
specific one of said output signals (c1, c2, c3 ...) is only extracted if said coherence
function (γ), said amplitude difference (amp) and said phase- or group delay difference
25   (phase, τ) all corresponds to the specific target intervals for that specific output signal
(c1, c2, c3...).

7. Method according to any of the preceding claims, <u>characterised in</u> that said extraction
of a given one of said output signals (c1, c2, c3 ...) is carried out on for instance fast
30   Fourier transforms of a given pair of input signals (L, R; L', R'; L", R"; ....), where said
given pair of input signals in the case of the first extracted output signal (c1) is the
original left and right channel signals (L, R), in the case of the second extracted output
signal (c2) is the first residual left and right channel signals (L', R'), in the case of the
third extracted output signal (c3) is the second residual left and right channel signals (L",
35   R") etc., where said fast Fourier transforms of a given pair of input signals are multiplied

by equal filter functions H(z) formed by said comparison of the determined coherence function (γ), the determined amplitude difference (amp) and the determined phase or group delay difference (phase, τ) with said target values hereof corresponding to the particular one of said output signals (c1, c2, c3 ...), which is to be extracted, and where

5 the multiplied versions of said fast Fourier transforms are inversely fast Fourier transformed (527) so that the two resulting time domain signals (535, 536), after individual appropriate scaling hereof, can finally be added (529) to form a first version (c1', c2', c3' ...) of that particular output signal (c1, c2, c3 ...) as a linear combination of said given pair of input signals, where said steps of fast Fourier transform, multiplication

10 and inverse fast Fourier transform are procedural steps of for instance the method known as FAST CONVOLUTION.

8. Method according to claim 7, <u>characterised in</u> that said output signals (c1, c2, c3, ...) are formed by amplification (530) followed by a post delay (531) of said first version (c1',

15 c2', c3' ...) of the output signals (c1, c2, c3 ...).

9. Method according to claim 7, <u>characterised in</u> that said filter function H(z) is a logical AND function, i.e. a function with output values of 1 or substantially 0, obtained by comparison at each frequency component of said coherence function, said amplitude

20 difference and said phase or group delay difference with corresponding target intervals corresponding to the particular one of said output signals to be derived, where H(z) is given by either the equation:

$$H(z) = (\gamma 1 < \gamma(z) < \gamma 2) \text{ AND } (amp1 < amp(z) < amp2) \text{ AND } (phase1 < phase(z) < phase2)$$

25 or by

$$H(z) = (\gamma 1 < \gamma(z) < \gamma 2) \text{ AND } (amp1 < amp(z) < amp2) \text{ AND } (group\ delay1 < group\ delay(z)$$
$$< group\ delay\ 2) \text{ AND } (-phase,max < phase(z) < +phase,max,\ where\ phase,max\ is\ less$$
than 180 degrees, preferably approximately 170 degrees..)

30 10. Method according to claim 7, <u>characterised in</u> that said filter function H(z) is a product at each frequency component of continuos functions of the values of the coherence function, the amplitude difference, the phase difference and/or the group delay difference, where the parameters of these functions are chosen according to sets of target intervals corresponding to the particular one of said output signals to be extracted.

35

11. Method according to claim 10, <u>characterised in</u> that said continuos functions are Gaussian functions (normal distribution density function) of the values of the square of the coherence function, the amplitude difference, the phase difference and/or the group delay difference, where the parameters of these Gaussian functions (normal distribution

5   density function) (means and variances) correspond to sets of target intervals corresponding to the particular one of said output signals to be extracted.

12. Method according to claim 9, 10 or 11, <u>characterised in</u> that said filter function $H(z)$ is formed as a product of a logical function $H1(z;p)$ with output values of 1 or

10   substantially 0 according to whether the parameters p, which may be the coherence function, the amplitude difference, the phase and/or group delay difference, belongs to the corresponding target intervals, and a function $H2(z;q)$ which is a product of continuous functions according to claim 10 or 11, where q denotes the remaining parameters not contained in said function H1.

15

13. Method according to any of the preceding claims, <u>characterised in</u> that the determination of said first residual left and right channel signals (L', R'), second residual left and right channel signals (L", R") etc. is carried out by subtracting (528) said two inversely fast Fourier transformed (527) signals (535, 536) respectively from delayed

20   (525) versions of left and right input signals (522, 523), which input signals (522, 523) in the case of the first output signal (c1) are the original left and right channel signals (L, R), in the case of the second output signal (c2) are the first residual left and right channel signal (L', R'), in the case of the third output signal (c3) are the second residual left and right channel signals (L", R") etc.

25

14. Method according to claim 2, <u>characterised in</u> that said degree of linear dependency between frequency components corresponding to a given output signal $(c_1, c_2, c_3...)$ is being evaluated based on the determination of the coherence function (γ), the amplitude difference (amp) and on the phase or group delay difference (phase or τ) between the

30   input signals (L, R; L', R'; L", R"...) to the corresponding processing block at each separate frequency component of the signals.

15. Method according to any of the preceding claims, <u>characterised in</u> that said electroacoustic transducers are loudspeakers.

35

16. Device for converting two original input signals L(t) and R(t) constituting the signals in the left and right channel of a stereophonic signal into N output signals corresponding to N output channels, where N > 2, characterised in that said device comprises means for extracting said output signals (c1, c2, c3 ...) based on the instantaneous degree of linear

5   dependency between signals elements in said two input signals and utilising sets of requirements concerning characteristic differences between said two input signals, said requirements being specific for each of said output signals (c1, c2, c3 ...).

17. Device according to claim 16, characterised in that said device comprises N-2 blocks

10 (32,33,34;76,77,78) each with two input signals, where each of said blocks extract one of said output signals (c1, c2, c3 ...), and where each of said blocks (32,33,34;76,77,78) furthermore provides two residual output signals (L', R'; L", R"; L"', R"' ....), which residual output signals do not contain those scaled versions of frequency components, which have been extracted as said output signals (c1, c2, c3...).

15

18. Device according to claim 17, characterised in that said blocks (32,33,34;76,77,78) are coupled in series after each other such that the first one of said blocks (32; 76) as input signals receives said original input signals L(t) and R(t), extracts a first one of said output signals (c1) and provides a first pair of said residual output signals (L', R'), and the

20 second one of said blocks (33; 77) as input signals receives said residual output signals (L', R'), extract a second one of said output signals (c2) and provides a second pair of residual output signals (L", R"), and the third one of said blocks (34; 78) as input signals receives said second pair of residual output signals (L", R"), extracts a third one of said output signals (c3) and provides a third pair of residual output signals (L"', R"'), etc., until

25 a maximum of N-2 output signals (c1, c2, c3 ...) have been extracted, and that the pair of final residual output signals (L"', R"'), which are left over after the extraction of the final one of said output signals (c3) are used as two separate output signals from said device.

19. Device according to any of the preceding claims 16 to 18, characterised in that said

30 degree of linear dependency between frequency components is being evaluated based on the determination of the coherence function (γ) of said original input signals L(t) and R(t) and the determination of the amplitude difference (amp) between said original input signals L(t) and R(t) and on the phase- or group delay difference (phase or τ) between said original input signals L(t) and R(t) at each separate frequency component of the

35 signals.

20. Devise according to any of the preceding claims 16 to 18, <u>characterised in</u> that said degree of linear dependency between frequency components in a particular analysis block (73, 74, 75) is being evaluated based on the determination of the coherence function (γ), the amplitude difference (amp) and on the phase or group delay difference

5   (phase or τ) between the input signals (L, R; L', R'; L", R" ...) to the corresponding processing block at each separate frequency component of the signals.

21. Device according to claim 19 or 20, <u>characterised in</u> that said device comprises means for determining said coherence function (γ), said amplitude difference (amp) and

10   said phase or group delay difference (phase or τ) based on calculated values of the autospectra $G_{LL}(f)$ and $G_{RR}(f)$ and on the crossspectrum $G_{LR}(f)$ according to the following equations:

15
$$\gamma(f) = \frac{|G_{LR}(f)|}{\sqrt{G_{LL}(f) \cdot G_{RR}(f)}} \tag{1}$$

20
$$amp(f) = \sqrt{\frac{G_{LL}(f)}{G_{RR}(f)}} \tag{2}$$

25
$$phase(f) = angle(G_{LR}(f)) \tag{3}$$

30
$$\tau(f) = \frac{-d(continous\_phase(f))}{2\pi \cdot df} \tag{5}$$

22. Device according to any of the preceding claims 16 to 21, <u>characterised in</u> that said sets of requirements concerning characteristic differences between said two input signals for each of said blocks (32, 33, 34; 76, 77, 78) comprise a target interval of said

35   coherence function (γ), said amplitude difference (amp) and said phase or group delay

difference (phase or τ), which target intervals are specific for that particular block, and which target intervals may be functions of frequency.

23. Device according to any of the preceding claims 16 to 22, characterised in that each
5 of said blocks (32, 33, 34; 76, 77, 78) comprise means for carrying out a comparison between said coherence function (γ), said amplitude difference (amp) and said phase or group delay difference (phase or τ) with the respective one of said target intervals, and means having the effect that a specific one of said output signals (c1, c2, c3 ...) is only extracted if said coherence function (γ), said amplitude difference (amp) and said phase
10 or group delay difference (phase or τ) all corresponds to the specific target intervals for that specific output signal (c1, c2, c3 ...).

24. Device according to any of the preceding claims 16 to 23, characterised in that each of said blocks (32, 33, 34; 76, 77, 78) performs the extraction of the specific output signal
15 (c1, c2, c3 ...) for that block by multiplication in appropriate multiplication means (526) the fast Fourier transformed input signals to that specific block with a filter function H(z), which filter function is the same for said two input signals to that particular block, which filter function H(z) is based on said comparison and thereafter providing said filtered input signals to inverse fast Fourier transform means (527) and thereby providing a pair
20 of signals (535, 536), which are provided to an addition means (529), the output signal of which is provided to a gain means (530) and thereafter to a delay means (531), the output signal of which is the desired output signal (c1, c2, c3 ...) of that particular block.

25. Device according to claim 24, characterised in that said filter function H(z) is provided
25 as the output signal from a logical AND means (517), this output signal taking on either the value 1 or substantially 0 according to the following expression:

$H(z) = (\gamma 1 < \gamma(z) < \gamma 2)$ AND (amp1 < amp(z) < amp2) AND (phase1 < phase(z) < phase2) or by

$H(z) = (\gamma 1 < \gamma(z) < \gamma 2)$ AND (amp1 < amp(z) < amp2) AND (group delay1 < group delay(z)
30 < group delay 2), AND ($-$phase,max < phase(z) < +phase,max, where phase,max is less than 180 degrees, preferably approximately 170 degrees..)

26. Device according to claim 24 characterised in that said filter function H(z) is a product at each frequency component of continuos functions of the values of the coherence
35 function, the amplitude difference, the phase difference and/or the group delay

difference, where the parameters of these functions are chosen according to sets of target values for said coherence function, said amplitude difference, said phase difference and/or group delay difference corresponding to the particular one of said output signals to be extracted.

5

27. Device according to claim 26, characterised in that said continuos functions are Gaussian functions (normal distribution density functions) of the values of the square of the coherence function, the amplitude difference, the phase difference and/or the group delay difference, where the parameters of these Gaussian functions (normal distribution density functions) (means and variances) correspond to sets of target values corresponding to the particular one of said output signals to be extracted.

10

28. Device according to claim 25, 26 or 27, characterised in that said filter function $H(z)$ is formed as a product of a logical function $H1(z;p)$ with output values of 1 or substantially 0 according to whether the parameters p, which may be the coherence function, the amplitude difference, the phase and/or group delay difference, belongs to the corresponding target intervals, and a function $H2(z;q)$ which is a product of continuous functions according to claim 10 or 11, where q denotes the remaining parameters not contained in said function H1.

15

20

29. Device according to any of the claims 16 to 20, characterised in that said residual output signals (L', R'; L", R"; L''', R'''; ...) in each of said blocks (32, 33, 34; 76, 77, 78) are obtained by subtraction in appropriate subtraction means (528) of said output signals (535, 536) provided from said inverse fast Fourier transform means (527), utilising said method of FAST CONVOLUTION, from the input signals to that particular block (32, 33, 34; 76, 77, 78) after these have been delayed in delay means (525) to compensate for the processing delay in said fast Fourier transform means (524) and in said inverse fast Fourier transform means (525).
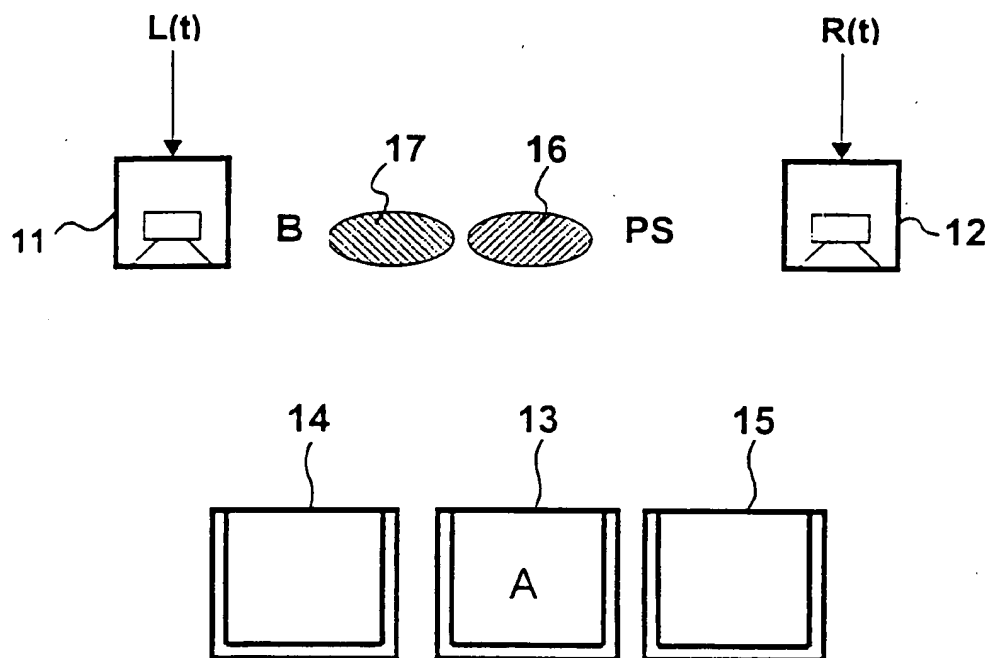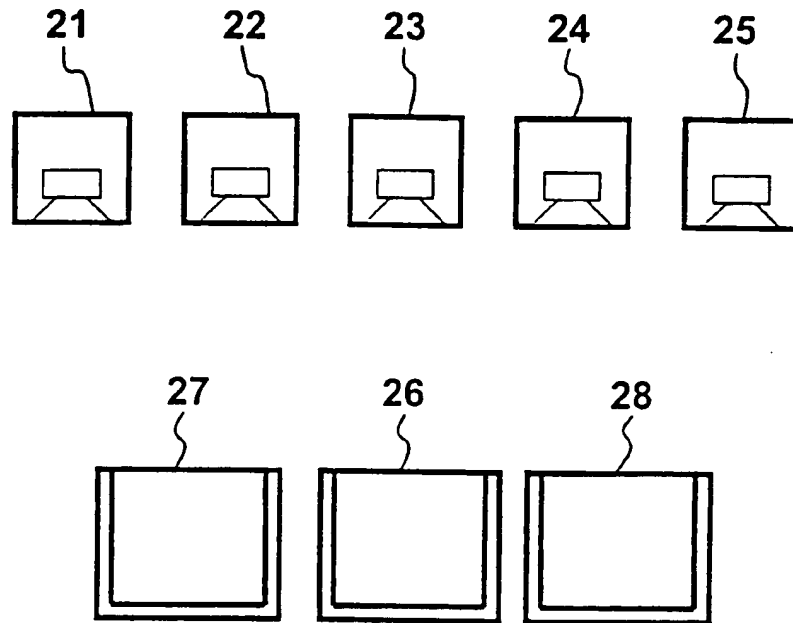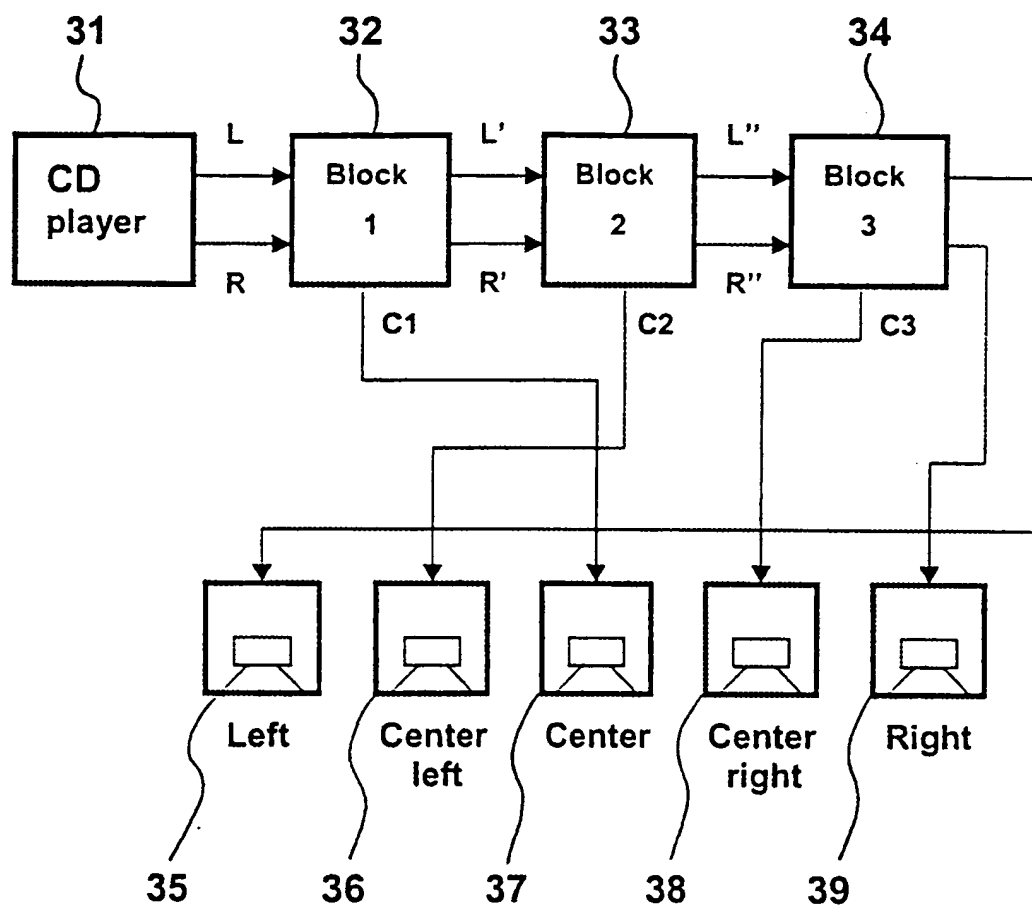
25

30

Fig. 1

Fig. 2

31	32	33	34

CD player →L→ Block 1 →L'→ Block 2 →L''→ Block 3

CD player →R→ Block 1 →R'→ Block 2 →R''→ Block 3

C1   C2   C3

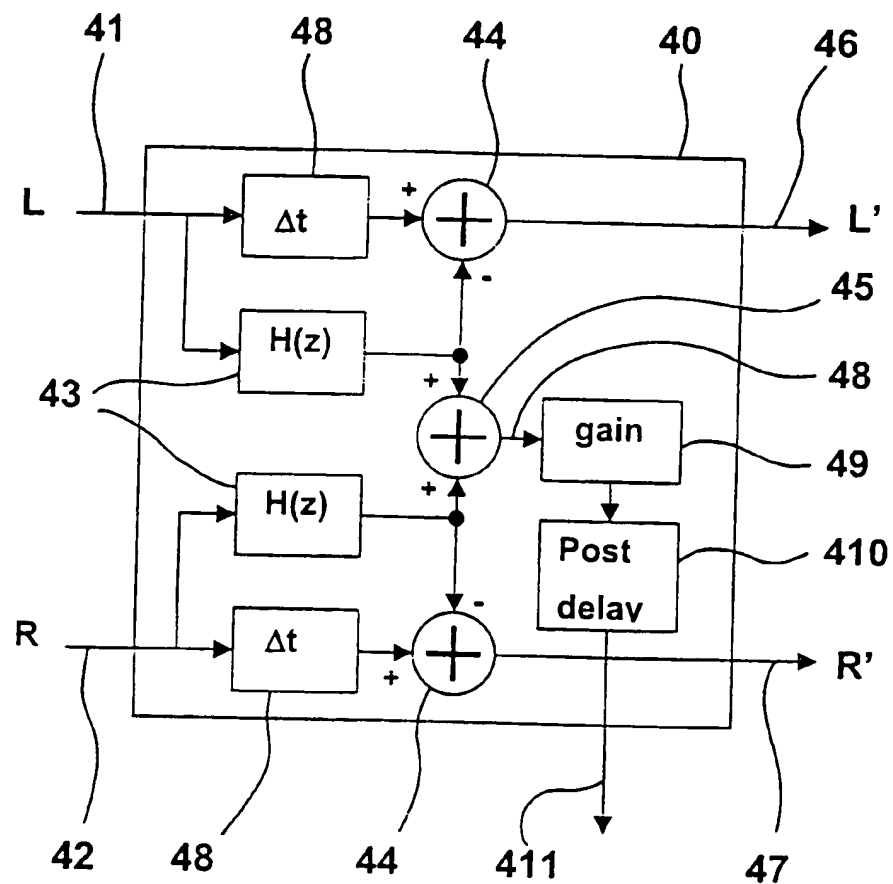Left   Center left   Center   Center right   Right
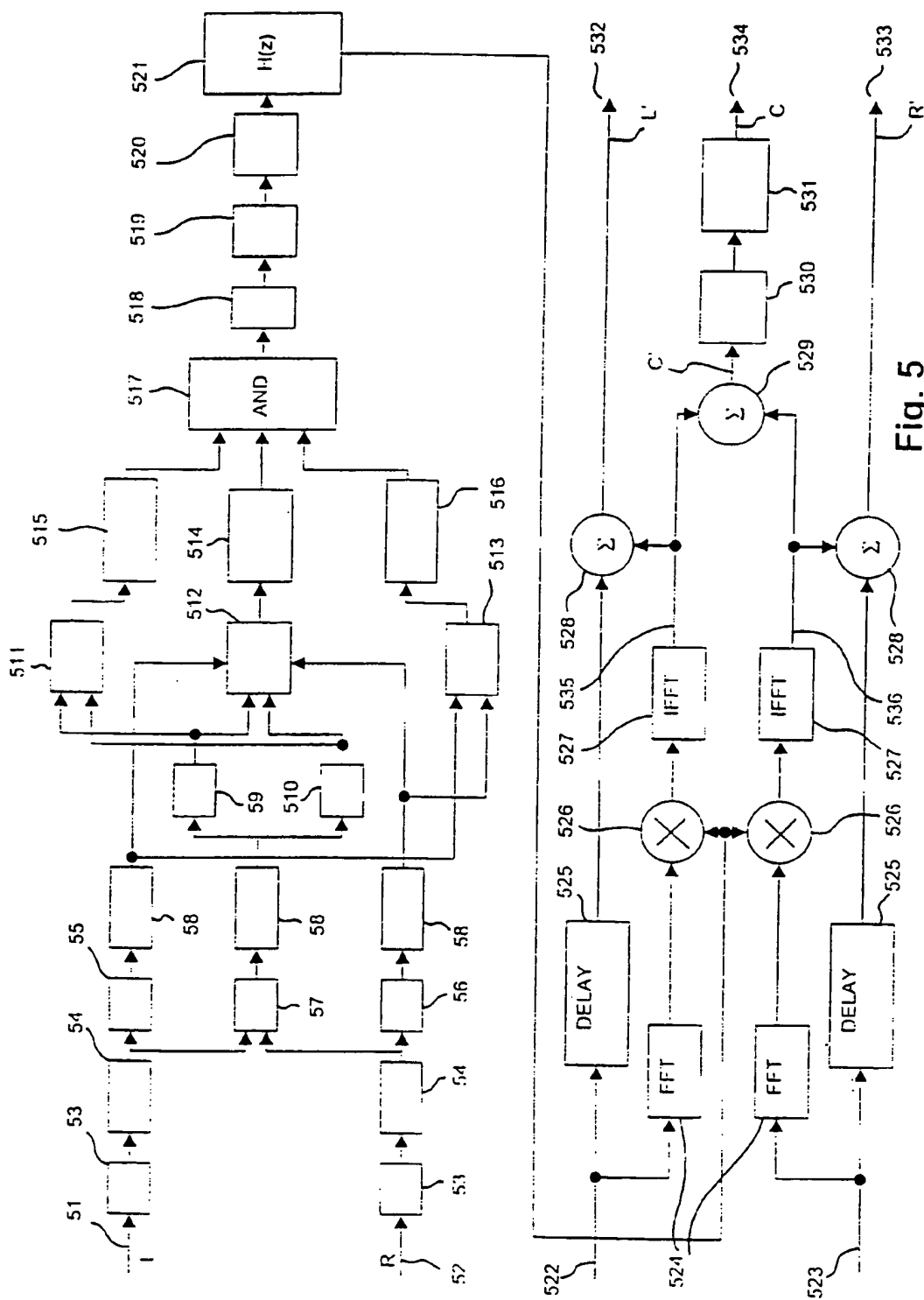
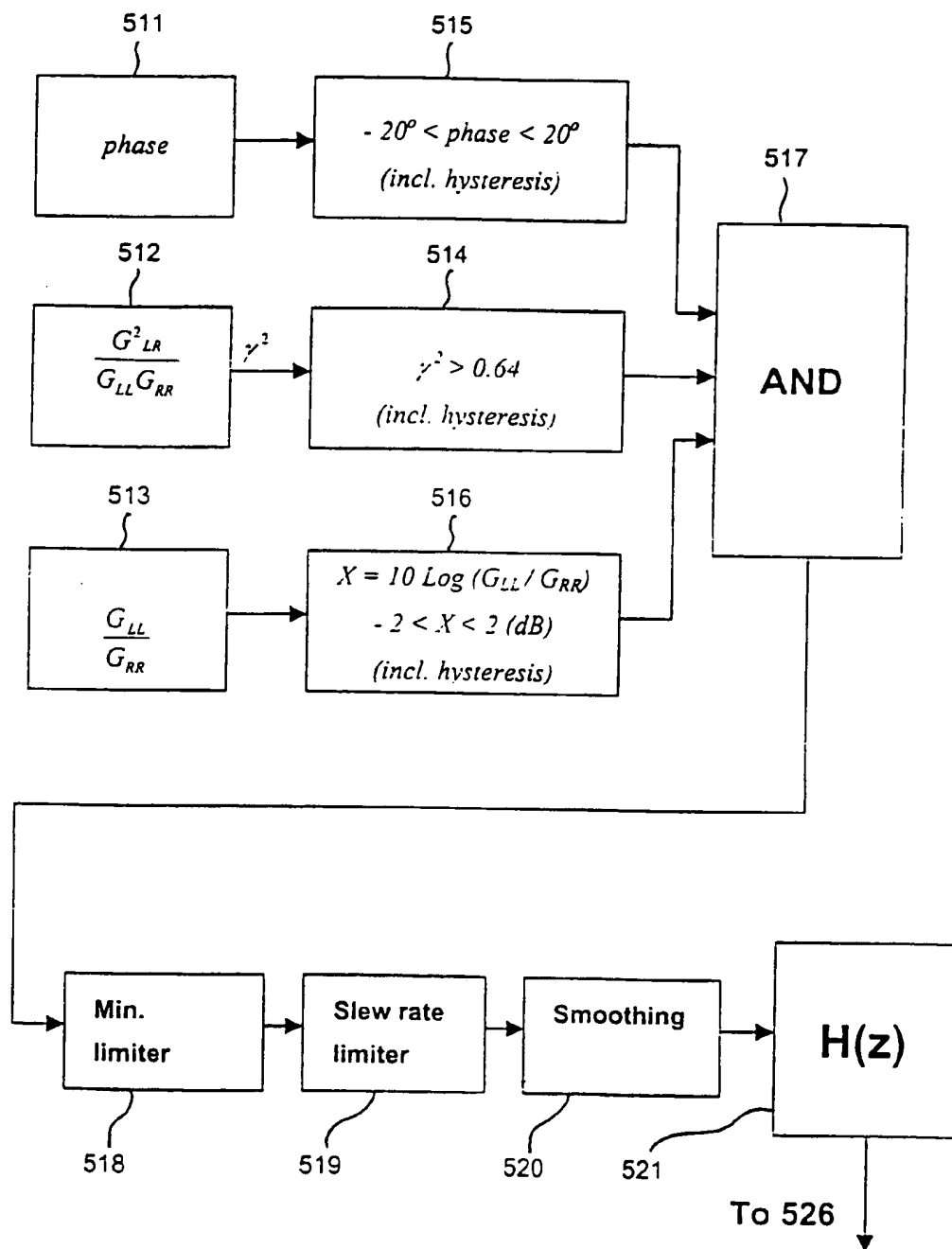35   36   37   38   39
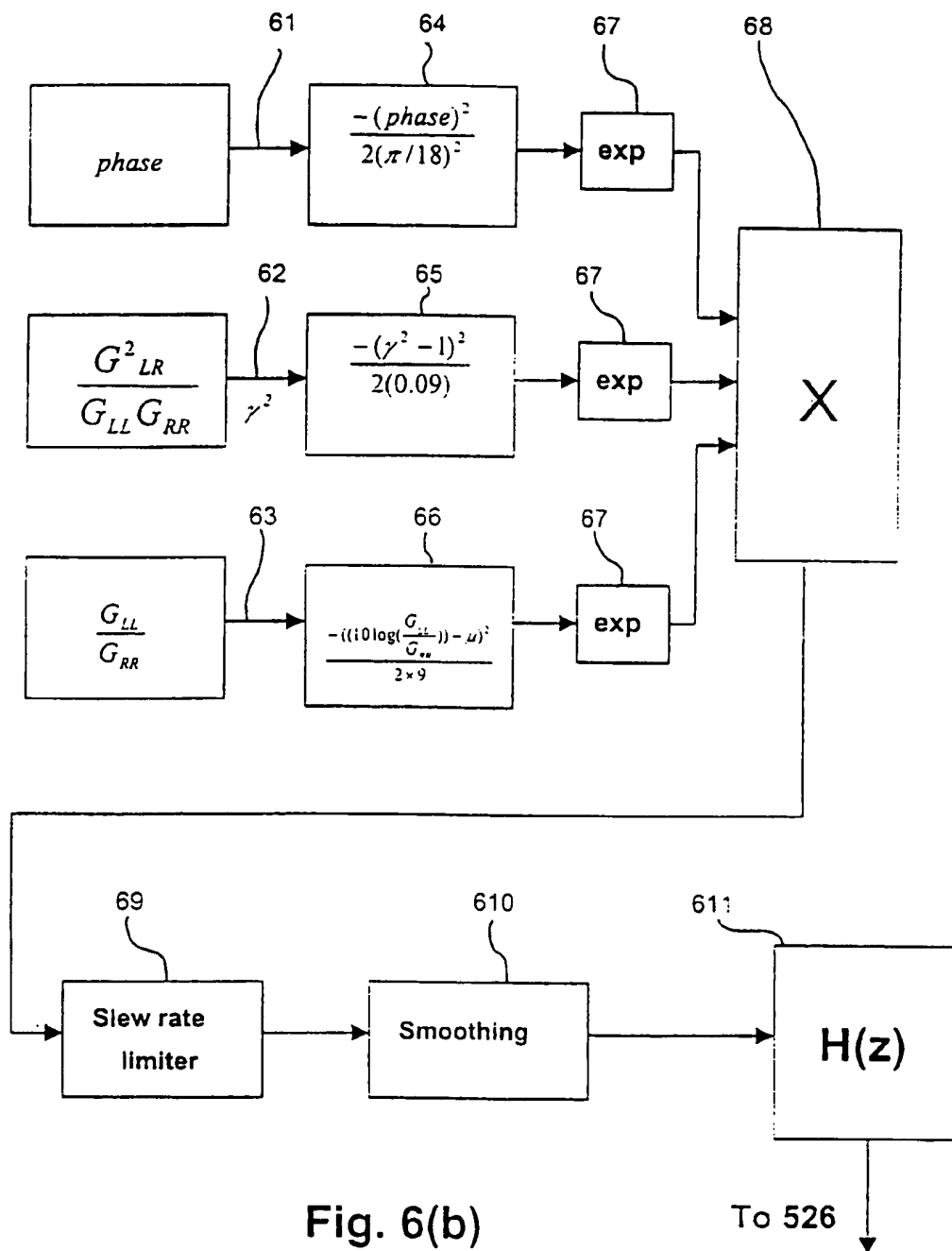
Fig. 3

Fig. 4

Fig. 5
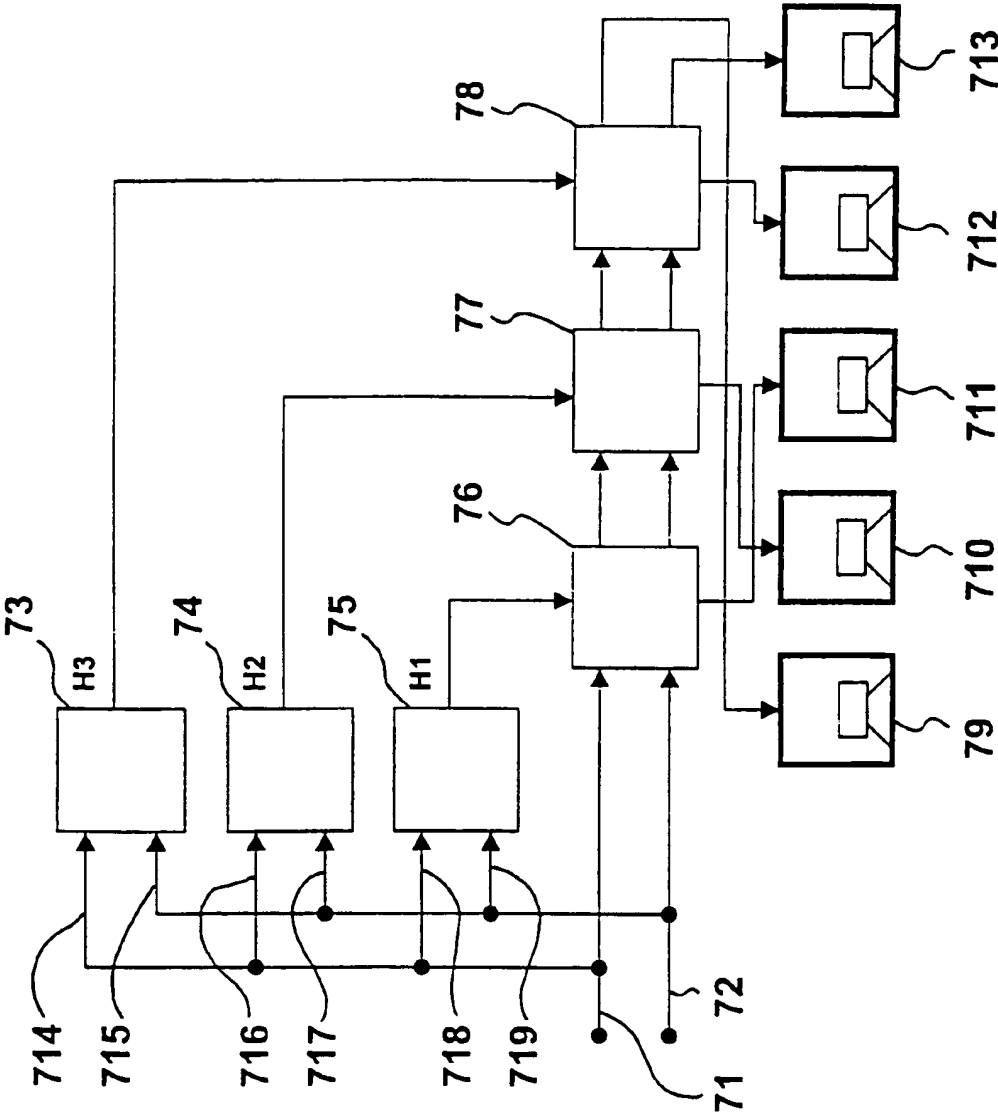
6/8

Fig. 6(a)

Fig. 6(b)

Fig. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 00/00068

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04S 3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H04S

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5870480 A (DAVID GRIESINGER), 9 February 1999 (09.02.99), column 3, line 60 - column 5, line 46, claim 6 --	1,16,19-21
A	US 5528694 A (LEON VAN DE KERKHOF ET AL), 18 June 1996 (18.06.96), column 1, line 53 - column 2, line 67; column 3, line 46 - column 4, line 8; column 6, line 4 - line 20 --	1,3-4
A	US 5046098 A (DOUGLAS MANDELL ET AL), 3 Sept 1991 (03.09.91), column 2, line 45 - column 3, line 28 --	1,3-4

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 October 2000	27. 11. 2000
Name and mailing address of the International Searching Authority European Patent Office P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel(+31-70)340-2040, Tx 31 651 epo nl, Fax(+31-70)340-3016	Authorized officer Erik Veillas/LR Telephone No.

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 00/00068

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	M. OMOLOGO et al:Acoustic event localization using a crosspower-spectrum phase based technique. Acoustics,Speech, and Signal Processing, 1994 IEEE International Conference on. Volume 2, pages II/273 - II/276. 19 - 22 April 1994. ISBN 0-7803-1775-0. See the whole document	1,3-4,16, 19-21
A	US 5136650 A (D. GRIESINGER), 4 August 1992 (04.08.92), see the whole document	1,16

US	5870480	A	09/02/99	AU	3665997	A	10/02/98
				CN	1228237	A	08/09/99
				EP	0923848	A	23/06/99
				US	5796844	A	18/08/98
				WO	9804100	A	29/01/98
				AU	5158998	A	29/05/98
				WO	9820708	A	14/05/98
US	5528694	A	18/06/96	DE	69423922	D	00/00/00
				EP	0608937	A,B	03/08/94
				JP	6253398	A	09/09/94
US	5046098	A	03/09/91	AU	594939	B	22/03/90
				AU	611337	B	06/06/91
				AU	4791490	A	10/05/90
				AU	5436586	A	11/09/86
				CA	1284112	A	14/05/91
				CN	1013822	B	04/09/91
				CN	86102314	A	22/10/86
				DE	3607610	A,C	18/09/86
				FR	2578707	A,B	12/09/86
				GB	2174275	A,B	29/10/86
				GB	2208780	A,B	12/04/89
				GB	8822178	D	00/00/00
				IT	1207068	B	17/05/89
				IT	8619656	D	00/00/00
				JP	2824642	B	11/11/98
				JP	61251400	A	08/11/86
				NL	192656	B,C	01/07/97
				NL	8600595	A	01/10/86
				US	4799260	A	17/01/89
				US	4941177	A	10/07/90
US	5136650	A	04/08/92	NONE			